

Bricolo：ブリコラージュ風発話生成システムの提案

Bricolo: Proposal for a Bricolage-like Utterance Generation System

天野 僚^{1*} 長谷川 孔明¹ 岡田 美智男¹
Ryo Amano¹, Komei Hasegawa¹ and Michio Okada¹

¹ 豊橋技術科学大学 情報・知能工学系

¹ Department of Computer Science and Engineering, Toyohashi University of Technology

Abstract: 日常の雑談では、相手からの反応を手掛りに断片的な思いをかき集め、言葉を並べていることが多い。このような発話スタイルを、筆者らは〈ブリコログ〉と呼んでおり、話し手と聞き手にレジリエントな関係性が構築できるものと考えている。本稿では、〈ブリコログ〉を生成するシステム〈Bricolo〉の基本コンセプトを述べ、その実装方法や効果について議論する。

1 はじめに

ロボットが私に何か伝えたいようだ。「隣町にできたんだって、隣町に、新しく...」(なにができたの?)「えーっと、ベーカリーらしいね。」(そうなんだ、なにがあるのかな?)「クリーム、なんかクリームが挟まった、パンで...」(うーんと、マリトッツォかな?)「そうそう、それ!それが人気らしいね。」... ロボットにしては少しぎこちないものの、ヒトと会話しているような温かみを感じ、なぜかもっと会話を続けたい。

このロボットがするような発話は私たちの日常会話で何気なく行われている。そのような日常会話では、話し手が頭の中で断片的な思いをかき集めながら話を紡ぎ、一文ごとに内容が完全に決まる前から発話が始まっていることが多い。一方で、聞き手は頭の中で断片的な発話を再構築しながらメッセージを受け取り、発話内容を解釈していると考えられる [1]。これを別の視点から捉えてみると、断片的な発話は聞き手の積極的参加を引き出す効果があるように思われる。

例えば、「明日この教室で、3時間目に数学のテストがあるよ。」と言われても、聞き手はただ何も考えることなく一方的にメッセージを受け取ってしまう。しかしながら、「テストが、数学のテストがあるよ。えっと、3時間目に、明日、あの一、この教室で。」というような断片的な発話であれば聞き手はメッセージの再構築を行うため、話し手と聞き手のインタラクティブな関係性が生じる。つまり、日常での会話は正しい文法で整えられた発話文でなくとも聞き手は内容を理解し、コミュニケーションがとれるうえ、断片的な発話の方が話し手に寄り添い合うことができると推察できる。

話し手は断片的な思いをかき集めるようにして言葉

を並べ、聞き手の様子によって臨機応変に発話が形成される。我々は発話形態のことを「ブリコログ (bricologue)」¹と呼んでいる。本稿では、「ブリコログ」を生成するシステム〈Bricolo (ブリコロ)〉のコンセプトやインタラクションについて述べ、その可能性を議論する。

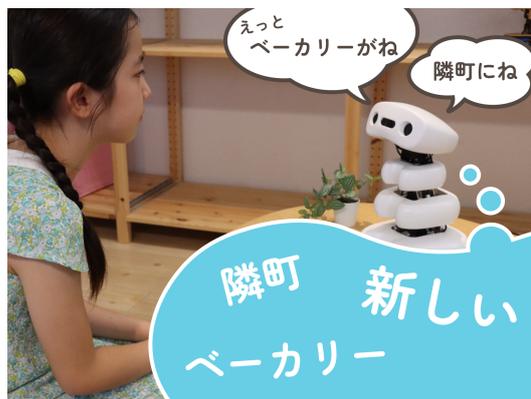


図 1: 〈Bricolo〉のコンセプトイメージ

2 研究背景

2.1 発話における漸次性・漸次的精緻化

私たちが日常で行う会話には「漸次性 (incrementality)」とよばれる基本的性質が含まれている [1]。漸次性とは次の2つの意味で用いられる。1つは「相手の知らない情報だけを伝える」という意味であり、もう1

*連絡先：豊橋技術科学大学 情報・知能工学系
〒441-8154 愛知県豊橋市天伯町雲雀ヶ丘 1-1
E-mail: amano.ryo.yt@tut.jp

¹ 「ブリコログ」は「ブリコラージュ (bricolage)」と「モノログ (monologue)」の合成語である。

つは「発話を構成していくとき、新たな情報を追加していくために発話の断片を思いつくままに次々に伝えていく」という意味で用いられる。前者は、お互いが共有する知識や情報については繰り返すことなく、そのとき相手がまだ知らないことだけを伝えるということを目指す。後者は、話し手が発話の断片を思いつくままに伝え、受け手側はその断片から発話を自ら再構築しながら、そのメッセージを解釈していくことを指す。

自らが表現したいイメージと実際に発話となって表現されたものとの間には「ずれ」が生じることもある。この「ずれ」は最後まで残るものだが、私たちは「ずれ」を小さくし精緻化していくためである。この精緻化の程度は与えられた時間によって変化する。この性質を「anytime性」といい、こうした性質は発話における「漸次的精緻化 (incremental elaboration)」の振る舞いと呼ばれている [1]。

したがって、日常での発話は必ずしも文として整った最適なものではなく、またそれを踏まえても十分にコミュニケーションがとれると考えられる。

2.2 ブリコラージュ

ブリコラージュとは、クロード・レヴィ＝ストロースが『野生の思考』の中で紹介した言葉であり [2][3]、「日曜大工・素人大工」という意味をもつフランス語であるが、ここでは「ありあわせをかき集めてその場を凌ぐこと」という意味で用いる。したがって、2.1節で述べたように日常での発話は「漸次性」を有するため、日常の発話はブリコラージュ風の発話であると捉えられる。

2.3 新たな発話形態〈ブリコローグ〉

コミュニケーションの形態を大きく分類すると2つに分けられる [5]。1つは、話し手の独り言や、話し手と聞き手の関係が固定化して一方の話し手が長い時間聞き手に向かって話し続ける〈モノローグ (monologue)〉であり、もう1つは話し手と聞き手が交互に発話しあってやり取りする〈ダイアローグ (dialogue)〉である。

本稿では、話し手が自らの伝えたい情報を断片的に表現し、その断片的な言葉を寄せ集め、聞き手の様子によって臨機応変に発話が形成される発話形態を提案する。ブリコラージュを由来とする新たな発話形態を〈ブリコローグ (bricologue)〉と名づけ、その発話生成システムの実現を狙っている。

2.4 会話における行為-知覚循環

漸次的精緻化の振る舞いが行われる会話では、話し手は聞き手の行為を伺いながら自己完結しない言葉足らずな発話をし、一方で聞き手は話し手の言葉足らずな部分を補いながら傾聴している。この行為と知覚の循環を「行為-知覚循環 (acting-perceiving cycle)²」とよぶ。この循環が破綻するとき、すなわち聞き手あるいは話し手の行為が放棄されるとき、会話は成立しなくなってしまう。このように「行為-知覚循環」は、行為するために知覚し、知覚するために行為する必要がある。

これまで松下らは、聞き手からの「あなたの話を聞いていますよ。」という聞き手性 (hearer-ship) をリソースとし、発話タイミングや内容を相手の様子を伺いながら発話文を生成する〈Talking-Ally〉の研究を行ってきた [4]。このロボットには、相手の姿を知覚することによって自分の行為を決定するインタラクションが備わっている。

本稿で提案する〈Bricolo〉は、〈Talking-Ally〉のもつインタラクションを活用した上に、ロボットがブリコローグで発話するようになっている。ヒトに会話への参加余地をさらに与えることで、より行為-知覚循環を強固にするものとなっている。

3 〈Bricolo〉の構成

3.1 発話生成系〈Bricolo〉

本システムは、ヒトの発話や傾聴の様子を手掛りにロボットの発話生成を行うもので、〈Bricolo (ブリコロ)〉と呼ぶ。

〈Bricolo〉は、ヒトとロボットとの行為-知覚循環と、ロボットによるブリコローグを実現している。ヒトからロボットに話しかけると、その内容に応じて断片的な言葉が漸次的に発話されていくことで、ロボットがブリコローグで発話する。一方、ロボットが発話している際のヒトの傾聴の様子によってロボットの話し方が変化する。

本研究で扱うロボットは、発話によって聞き手との多様な関わりを引き出すソーシャルロボット〈Talking-Bones〉を用いる。〈Talking-Bones〉の外観を図2に示す。

²双方の行為を主体としているため、ここでは「知覚-行為循環」ではなく「行為-知覚循環」と表記する。



図 2: 〈Talking-Bones〉の外観

3.2 システムデザイン

〈Bricolo〉のシステム構成を図3に示す。〈Talking-Bones〉ではヒトの音声と顔情報を取得する。

音声認識データはテキスト生成AIの入力とし、AIの出力を形態素解析を介して断片的な言葉を生成する。〈Talking-Bones〉は生成される言葉のテキスト情報を受け取りスピーカによって読み上げ、モータによって動きの振る舞いを実装している。

一方、ヒトの顔情報データについては、ヒトの顔が〈Talking-Bones〉に向いているか否かによって、〈Talking-Bones〉の発話に変化が起きるように実装している。例えば、ヒトが興味をもつように話を聞いている（顔がこちらに向いている）ときは揚々と話すが、興味をもっていない（顔がこちらに向いていない）ようであれば、「えっと〜」や「あの〜」といったフィラーが増え、辿々しくなる。

〈Bricolo〉のアーキテクチャは、BehaviorレイヤとGenerativeAIレイヤの2レイヤ構造となっている（図4）。

Behaviorレイヤは、ロボットの発話やモータ制御などの振る舞いを管理する。一方で、GenerativeAIレイヤはテキスト生成AIを導入しており、ロボットの発話内容を生成する。これらのレイヤは双方向でテキスト情報や顔認識情報について通信を行う。

Behaviorレイヤで音声認識でヒトが話す内容やヒトの顔情報を把握し、その内容を手掛りとして、GenerativeAIレイヤでロボットの発話内容を生成させる。

生成した発話内容は、Behaviorレイヤによってロボットが実際に発話する。

〈Bricolo〉はRobot Operating System 2（通称ROS2）によって構成される。

3.3 ハードウェアデザイン

本研究で扱うプラットフォームである〈Talking-Bones〉は骨を模したデザインとなっており、卓上に配

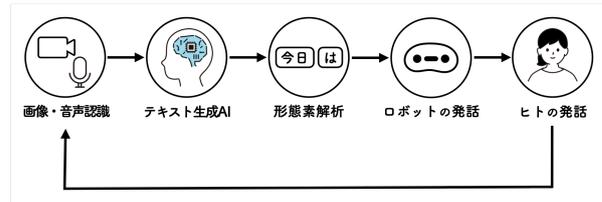


図 3: 〈Bricolo〉のシステム構成

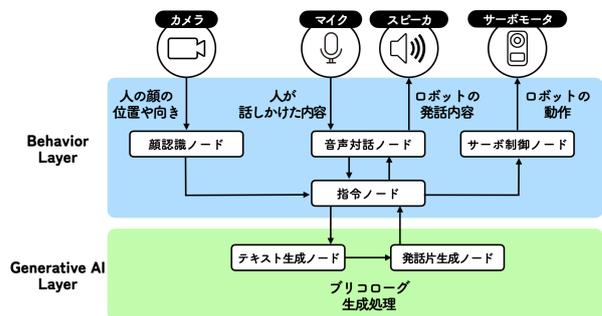


図 4: 〈Bricolo〉のソフトウェア構成

置できる大きさのソーシャルロボットである [6]。本システムでは、プリコログで話すエージェントとして活用する。

〈Talking-Bones〉のハードウェア構成を図5に示す。顔認識用のカメラと発話用のスピーカ、音声認識用のマイクから構成されている。内部にはサーボモータを3つ搭載しており、頭部と胴体の振る舞いをコントロールしている。また、胴体下部の接合部にバネを用いていることで、生き物らしさを創出している。

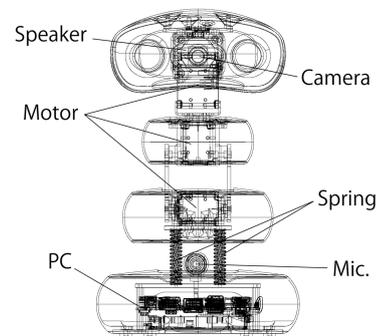


図 5: 〈Talking-Bones〉の構成

4 インタラクシオンデザイン

〈Bricolo〉のインタラクシオンは、ロボットとの会話を中心となり行われる。そのインタラクシオンの過程は以下の通りである。

1. ヒトがロボットに話しかける。
2. 話しかけた内容に応じて、ロボットが応答内容を生成する。
3. ロボットがヒトの顔を認識しているとき、ロボットが断片的に発話する。
4. ロボットがヒトの顔を認識しなくなったとき、フィラーが増えロボットの発話がたどたどしくなる。

上記の過程により、図6で示すインタラクシオン例のようにロボットは相手の様子を伺いつつ言葉の順序を完全に決める前から発話を始め、漸次的発話を行う。

また、過程の中で発話されないブランクが生じる。このブランクでは、適宜「えっと」や「あのー」といったフィラーを発話させる。

この過程により、ロボットはブリコローグをすることになる。よって、ヒトはロボットから聞き手が引き出されつつ、ロボットの発話を再構築しながら受け取るというインタラクシオンが行われ、ヒトがロボットとの会話に自然と埋め込まれる効果が期待できる。

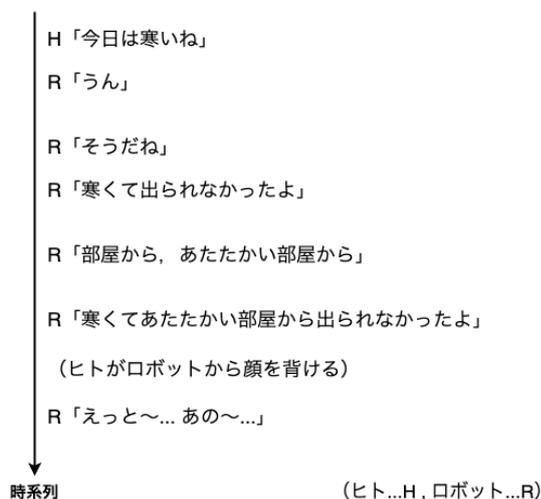


図6: 〈Bricolo〉のインタラクシオン例

5 最後に

本研究では〈ブリコローグ〉という発話スタイルや〈ブリコローグ〉生成システムである〈Bricolo〉のコン

セプトとインタラクシオンをデザインし、その可能性について紹介した。

また、このシステムによって、ロボットとの会話における行為-知覚循環が生まれ、より身近なロボットに感じられる効果を期待する。

今後はシステム面の改良を実施し、ロボットによるブリコローグらしさを向上させ、期待する効果についての実証実験を実施していきたい。

謝辞

本研究の一部は、愛知県が公益財団法人科学技術交流財団に委託し実施している「知の拠点あいち重点研究プロジェクト第IV期(第4次産業革命をもたらすデジタル・トランスメーション(DX)の加速)」により行われた。ここに記して感謝の意を示す。

参考文献

- [1] 岡田美智男: 『口ごもるコンピュータ』; 共立出版, (1995).
- [2] クロード・レヴィ=ストロース(著):大橋(訳): 『野生の思考』; みすず書房, (1976).
- [3] 小山尚之: 現代社会の日常生活における野生の思考: 未開人たちのもたらしたものと A. ブルトンに関する-考察; 東京商船大学研究報告, 人文科学, 51 巻, pp.75-88, (2000).
- [4] 松下仁美, 香川真人, 山村祐之, 岡田美智男: 非流暢性を伴うロボット 〈Talking-Ally〉 の発話調整方略とその聞き手に対する適応に関する研究; ヒューマンインタフェース学会論文誌, Vol.20, No.2, pp.255-268, (2018).
- [5] 田島充士: 教育実践を理解するためのバフチン・ダイアログ論 豊かな異文化交流の実現; 言語文化教育研究学会, 第16巻, pp.260-278, (2018).
- [6] 小野田慎平, 西脇裕作, 窪田裕大, 大島直樹, 岡田美智男: 子どもたちはときどきモノ忘れするロボット 〈Talking-Bones〉 とどのように関わるのか?; ヒューマンインタフェース学会論文誌, Vol.23, No.2, pp.213-226, (2021).